CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

761177Orig1s000

STATISTICAL REVIEW(S)

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #:	BLA 761177
Drug Name:	Lonapegsomatropin
Indication(s):	Pediatric growth hormone deficiency
Applicant:	Ascendis
Date(s):	Submit Date: 6/25/2020
	Review Due Date: 2/25/2021
	PDUFA Date: 6/24/2021
Review Priority:	Standard
Biometrics Division:	II
Statistical Reviewer:	Alexander Cambon, PhD
Concurring Reviewers:	Feng Li, PhD
	Mark Rothmann, PhD, Division Director
Medical Division:	CDER/OND/OCHEN/DGE
Clinical Team:	Shivangi Vachhani, M.D.
	Marina Zemskova, M.D., Team Leader
Project Manager:	Sejal Kiani

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 EXECUTIVE SUMMARY

Ascendis Pharma is seeking approval for efficacy and safety of Lonapegsomatropin, a long acting growth hormone, delivered by weekly injection, for treatment of (b) (4) (b) (4)

Brief Overview of Clinical Studies

This submission encompasses one efficacy trial CT-301, an extension trial CT-301EXT, and a safety trial CT-302. This review focuses on the results from the efficacy trial. Study CT-301 was a 52-week, randomized, parallel active-controlled and open label design comparing Lonapegsomatropin administered by weekly injections, to Genotropin administered using daily injections.

In Study CT-301, the primary efficacy endpoint was Annualized Height Velocity (AHV) at week 52. The primary endpoint was compared between Lonapegsomatropin and Genotropin with a pre-specified non-inferiority margin of 2 centimeters (cm) per year, followed by a test of superiority if non-inferiority was established. The treatment difference was 0.80 centimeters per year (cm/yr) with 95% confidence intervals of (0.13, 1.47), which achieves statistical significance for superiority over Genotropin.

Major Statistical Issues

The efficacy study CT-301 achieved statistical superiority compared to Genotropin using the sponsor's pre-defined analysis of covariance (ANCOVA) analysis. However, the treatment difference is small and it is not clear that statistical superiority would be replicated in another study. This study on its own may not provide sufficient evidence for a determination of superiority to another approved drug.

1.1 Conclusion and Recommendations

For the primary analysis, Study CT-301 demonstrated non-inferiority of the study drug for the primary endpoint, Annualized Height Velocity (AHV), in comparison to Genotropin, a drug already approved for treatment of prepubertal children with GHD. The results also showed that the AHV in the Lonapegsomatropin group was greater than that of the Genotropin group with statistical significance. However, the magnitude of the treatment difference is small (b) (4)

2 INTRODUCTION

2.1 Overview

Lonapegsomatropin is a long-acting human growth hormone (hGH), designed to maintain the same mode of action and distribution in the body as daily somatropin, but with a once-weekly injection.

Previous Communications

Agency End of phase 2 (EOP2) statistics related comments to the sponsor (meeting date April 26, 2016) included the following:

- A multiple testing procedure that controls Type 1 error should be used for labeling claims to be considered for more than one endpoint.
- The proposed MMRM (Mixed Model Repeated Measures) does not adequately take into account missing data due to treatment discontinuation.
- Use of MMRM may lead to a test statistic that does not have a standard normal distribution if variances are unequal
- All observed data, including post-discontinuation data, should be included in analysis. Detailed efforts should be undertaken to prevent missing data.

Statistical comments for Type B Meeting Request were communicated to the sponsor on August 25, 2016:

> Please provide a justification for the choice of the non-inferiority margin that is based on the effect of Genotropin on annual height velocity at 12 months from previous clinical trials in the same or similar clinical trial setting as the proposed clinical trial. The non-inferiority margin should be informed by the considerations described in the draft Guidance for Industry – Non-inferiority Clinical Trials.
>
> We are interested in estimating the treatment effect based on the intent-to-treat (de facto) estimand, which considers the actual measurements of subjects regardless of adherence to treatment or use of subsequent therapy. The primary analysis should account for missing data in the primary endpoint in a fashion consistent with what the measurement would have been, had it been measured and should address missing data based on that information most relevant to what the measurement would have been had it been measured. Your proposed MMRM also does not appropriately address missing data as it treats the behavior of missing data for those patients who are off-treatment to be the same as that of observed data for those patients who are on-treatment in the same treatment arm. We would recommend addressing missing data on the primary endpoint by having the missing data from subjects who do not adhere to therapy represented by the data from those subjects on the same arm that also did not adhere to therapy but had the measurement for the primary endpoint.

The sponsor submitted responses to these comments on 9/13/2016, including the requested justification of the non-inferiority margin based on the draft guidance. Statistical comments submitted into DARRTS on 10/24/2016 included the comment:

Your justification for the non-inferiority margin [2 cm/yr] is adequate.

The Statistical Analysis Plan (SAP) Review submitted into DAARTS on January 2019 reiterated that the proposed MMRM model, which the sponsor proposed using in the event that there are not enough retrieved dropouts, still does not adequately address missing data due to treatment discontinuation. The following comments were also provided:

In your SAP Section 8.6, the testing hierarchy contains an endpoint (b) (4) (b) (4). However, this endpoint should not be considered as a key secondary endpoint. Please remove this endpoint from the testing hierarchy and update your statistical testing procedure as appropriate. We remind you that the significance level should be appropriately conserved in the testing hierarchy. In addition, provide a detailed test plan for each endpoint in the testing hierarchy.

We recommend that the primary efficacy analysis population consist of all randomized subjects who receive at least one dose of study drug.

Your SAP Section 8 says "Noninferiority tests will be based on a one-sided significance level of 0.025." However, the SAP Section 8.6 says "The familywise type-1 error of the study is controlled at alpha=(b) (4)", and your proposed testing procedure says that the test for superiority of the primary efficacy endpoint will be done "with alpha=(b) (4)". Please clarify this discrepancy.

We remind you that the adequacy of a single trial to support approval will be determined by its ability to support the efficacy claim based on the strength of the results. If only one clinical trial is conducted, then internal consistency across study subsets, evidence of an effect on multiple endpoints, and statistically very persuasive efficacy results will be considered in the evaluation. For additional information, refer to the following guidance for
industry:

https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm072008.pdf.

Statistical comments submitted into DARRTS on 3/25/2019 reiterated that an ANCOVA model be used for both primary and secondary endpoints for which labeling claims are planned, and that the imputation method should still not rely on a missing at random (MAR) assumption for patients who discontinue early, even though the missing rate is low and not likely to have a major impact on the results.

The Type B Pre-BLA statistics related comments (Meeting Date December 10, 2019) included the following:

1) Clinical Comment

We note that study CT-301EXT is still ongoing and you plan to include the results of interim analysis in this BLA. Thus, the adequacy of the data submitted to demonstrate the safety, purity, and potency of your proposed biological product for the use for which you

are seeking licensure will be a review issue. Please clarify how many patients in the CT-301EXT study are expected to reach final height by the time of the BLA submission.

2) Post-Meeting Comment:

We recommend that you submit to your IND a justification for the extrapolation of the short-term growth data (2 years) to final height before your BLA submission.

3) ISS and ISE Comments

Concerning safety analysis: In general, we agree with the overall ISS [Integrated Summary of Safety] proposal. However, we do not agree with the pooling strategy II for safety data. We do not agree that it is appropriate to combine the data from studies CT-301, CT-302 and CT-301EXT. These studies had different study design and different randomization ratio (i.e., single arm [studies CT-302 and CT-301EXT], vs. active controlled study CT-301), and were conducted in a different patient population (studies CT-302 and CT-301EXT in previously treated patients vs. study CT-301 in treatment-naive patients). Thus, a simple pooling of data from these studies would not be informative and can be misleading. The Summary of Clinical Safety (SCS)/Integrated Summary of Safety (ISS) should include separate summaries of safety for the individual Phase 3 studies (study CT-301, study CT-302, and study CT-301EXT) which should include the prespecified endpoints and subgroup analyses as outlined in the ISS SAP. In addition, we request you include in your BLA tabulations and narratives for all deaths, serious adverse events, and adverse events leading to discontinuation from the pediatric clinical development program. We may request additional data if safety concerns are identified during the review for which additional data may be informative.

Concerning efficacy analysis: Overall, we agree with the proposed presentation of efficacy results, i.e., including individual efficacy results from study CT-301 that provide pivotal efficacy data, and from Phase 3 CT-302 and Phase 3 CT-301EXT studies that provide supportive efficacy data for the proposed indication. Please note, our assessment of efficacy will focus on the collective evidence from individual studies. While we agree you may include additional pooled analyses sets assessing long-term efficacy of ACP-011 as supportive efficacy data, the interpretability of these analyses will be complicated, due to differences in studies design, previous exposure to different somatropin agents, differences in the population being studied, etc.

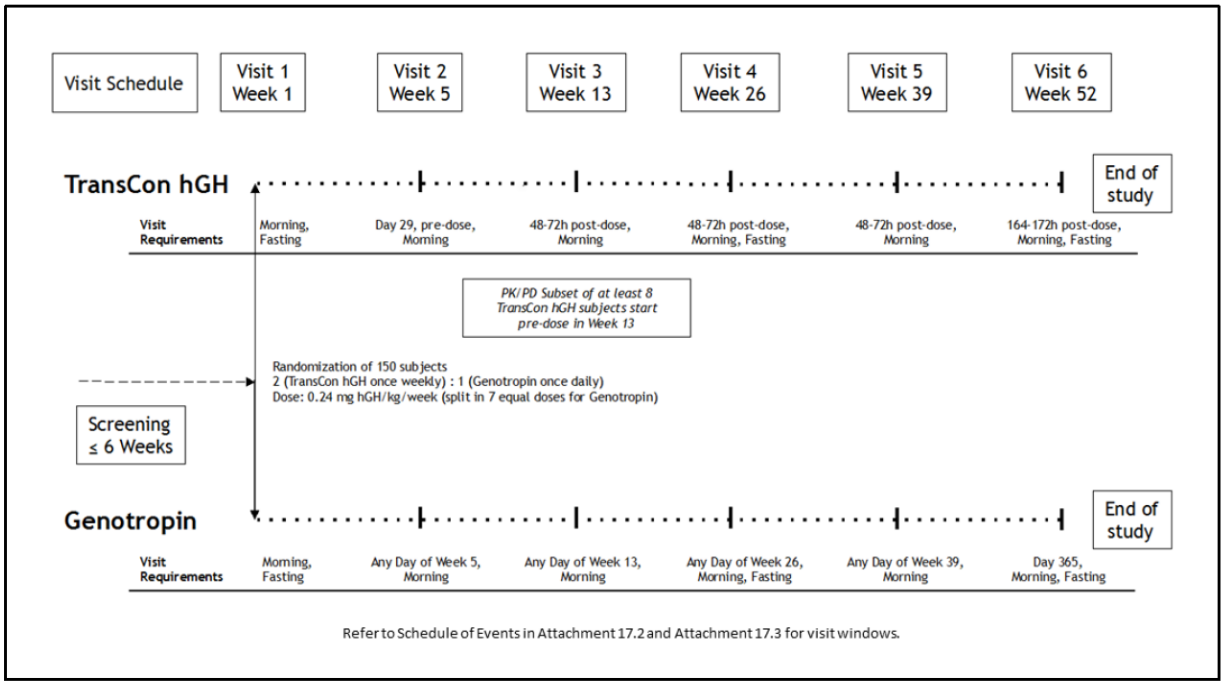
The FDA granted orphan drug designation on April 14, 2020.

This submission included one confirmatory 52-Week efficacy Study CT-301 (Figure 1), a one-arm extension Study CT-301EXT, and a one-arm 26-Week Study, CT-302.

2.2 Data Sources

The data and final study report for BLA 761177 were submitted electronically as an eCTD submission. The submission is archived at the following link.
\\CDSESUB1\evsprod\BLA761177\0001

Study CT-301 is a 52-Week phase 3 open-label randomized study comparing Lonapegsomatropin, administered using weekly injections, to Genotropin, administered using daily injections.



h = hour; hGH = human growth hormone; kg = kilogram; mg = milligram; PD = pharmacodynamics; PK = pharmacokinetic

Figure 1: Design for Study CT-301
Source – protocol for Study CT-301, Figure 5-2; TransCon hGH - Lonapegsomatropin

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The SDTM and ADaM data sets are located in the proper sections of the submission, and analysis reviewer guides are provided which define variables and their locations. I also checked for data quality issues and found the data quality to be satisfactory. I was able to replicate the sponsor's analyses for primary and secondary endpoints included in proposed label for Study CT-301.

Of note, the sponsor's efficacy dataset (ADEFF) for Study CT-301 includes 100 multiple imputations for each subject. I identified missing data for the primary endpoint in part by identifying the two patients who had different imputation results for each of the 100 imputations. The ADY (Analysis Relative Day) variable in the ADEFF dataset, was set at 365 for both patients, though the "Completer's Flag" variable was set at "N" only for these patients, and the "Analysis Datetime" variable was missing only for these two patients. It would seem that, since

the final assessments are missing for these patients, the ADY variable for these final assessments should be set to missing.

3.2 Evaluation of Efficacy

3.2.1 Study Design and Endpoints

The primary and secondary endpoints for Study CT-301 are shown in Table 1 below. These endpoints are assessed at 52 weeks. A hierarchical testing approach was only pre-specified for the first two endpoints in the protocol before the start of the study. There were no endpoints specified for safety study CT-302 or for extension study CT-301EXT, though descriptive results are given for AHV and change in height SDS (Standard Deviation Score).

Table 1: Primary and Secondary Endpoints – Study CT-301

Endpoint Type	Description
Primary	Annualized Height Velocity at 52 Weeks (Non-Inferiority*)
Secondary	Annualized Height Velocity at 52 Weeks (Superiority)
Secondary	Change in Height SDS at 52 Weeks

Abbreviations: SDS-Standard Deviation Score; *Non-inferiority margin-2 cm/year

3.2.2 Statistical Methodologies

3.2.2.1 Sponsor Approach

The sponsor's primary analysis population is the Intent to Treat (ITT) population consisting of all randomized treated patients. The sponsor's preferred estimand is the treatment policy estimand. The sponsor's defined primary analysis approach for continuous endpoints, including the primary endpoint of Annualized Height Velocity is ANCOVA. Treatment group and gender are included as factors in the model. Baseline age, peak growth hormone levels (log transformed) at stimulation test, and baseline height SDS minus average SDS of parental height are included as continuous covariates. Stratification factors include age (>3 to ≤6 and >6 years), peak GH levels in stimulation tests (≤5 ng/mL vs. >5 ng/mL), and gender. Missing final assessments were multiply imputed using the MAR assumption.

3.2.2.2 Reviewer Approach

Estimand

The preferred estimand is the treatment policy estimand which includes all data within the final 52-Week assessment window regardless of intercurrent events such as treatment discontinuation or initiation of alternative therapy. This is also the sponsor's pre-specified estimand. I also agree with the sponsor's ITT population (Section 3.2.2.1).

There were only two patients with missing final assessments, one on the Genotropin arm and one on the Lonapegsomatropin arm. The Genotropin patient with a missing final assessment was still on treatment, while the patient on Lonapegsomatropin discontinued treatment at 34 weeks. This was taken into account in my analysis. Since the missing data rate was so low (less than 2% on the Genotropin arm and less than 1% on the Lonapegsomatropin arm) I used a single imputation ANCOVA, with the fitted final assessment value from the ANCOVA model for the Genotropin patient, and the NMAR (not missing at random) assumption for the patient on Lonapegsomatropin. I used the patient's baseline AHV value to impute the missing final AHV assessment for the Lonapegsomatropin patient.

3.2.3 Patient Disposition, Demographic and Baseline Characteristics

A total of 358 patients were screened for study CT-301. Of these patients, 162 were randomized to Lonapegsomatropin (n=106) or Genotropin (n=56), and 105 patients were treated with Lonapegsomatropin compared to 56 patients on Genotropin. At the 52-week final assessment window, 104 patients on Lonapegsomatropin and 56 patients on Genotropin were still on treatment (Table 2). One patient on the Lonapegsomatropin arm discontinued treatment at 34 weeks and did not have a final assessment. This patient was withdrawn from the study by the parent/guardian. Another patient on the Genotropin arm completed 52 weeks of treatment but did not have a 52-week final assessment. No TEAE (treatment emergent adverse event) led to treatment discontinuation or death. There were also no reported deaths in this study.

Table 2: Descriptive statistics for patients having primary efficacy data, and patients discontinuing treatment.

Treatment	Patients Rand. / Treated	Patients Rand. Treated With BL	Disc. Treatment Early*	Disc. Treatment Early, and Missing	Did not Disc. Treatment Early, Missing	Missing	% Missing
Lonapeg.	105	105	1	1	0	1	1.0
Genotropin	56	56	0	0	1**	1	1.8

Abbreviations: Lonapeg- Lonapegsomatropin; Rand-randomized; BL-baseline measure; Disc-Discontinued; *Based on discontinuing 8 or more weeks early. Patient (b) (6), randomized to Lonapegsomatropin, discontinued treatment at 238 days after treatment start date and was "withdrawn by parent/guardian". **Patient (b) (6), randomized to Genotropin, completed treatment but did not have a final assessment.

The distributions of baseline demographic characteristics for Study CT-301 are shown in Table 3. Most patients identified as White race (152 of 161 patients, or 94%). This has implications for subgroup analysis, as there are not enough patients from other race subgroups to conduct meaningful subgroup analyses for race. For example, there were three Black/African Americans and one Asian. Patients were mostly from the regions of Europe and North America (139 of the 161 patients, or 86%), and 132 (82%) of the patients were male. Characteristics, including baseline height velocity and Height SDS seem evenly distributed between treatments arms. On average, patients had a baseline Height SDS of -3, and all patients had a baseline Height SDS of less than -1.

Table 3: Demographics and Baseline Characteristics by Treatment Arm - Study CT-301

Treatment Group	**Lonapeg.**	**Genotropin**
N per Group	105	56
Sex, n (%)		
Female (%)	19 (18)	10 (18)
Male (%)	86 (82)	46 (82)
Race, n (%)		
Asian	1 (1)	0 (0)
Black/AA	2 (2)	1 (2)
Multiple	0 (0)	1 (2)
White	100 (95)	52 (93)
Filipino	1 (1)	0 (0)
Half-Italian And Half-Japanese	1 (1)	0 (0)
New Zealand Maori (%)	0 (0)	1 (2)
Roma (%)	0 (0)	1 (2)
Age		
Mean (SD)	8.4 (2.7)	8.4 (2.8)
Median (min - max)	8.6 (3.2 - 13.0)	8.6 (3.2 - 12.8)
Age Group*, **n (%)**		
Age < 6 Years	25 (24)	14 (25)
Age >= 6 Years	80 (76)	42 (75)
Region, n (%)		
Europe	66 (63)	31 (55)
Middle East/ North Africa	6 (6)	8 (14)
North America	27 (26)	15 (27)
Oceania	6 (6)	2 (4)

Treatment Group	Lonapeg.	Genotropin
Ethnicity, n (%)		
Hispanic or Latino	5 (5)	2 (4)
Baseline BMI (kg/m)		
Mean (SD)	16.1 (1.8)	16.5 (2.2)
Median (min - max)	15.7 (13.2 - 22.2)	16.1 (13.7 - 24.7)
Baseline Height (cm)		
Mean (SD)	113 (14.1)	112 (15.3)
Median (min - max)	114 (87 - 139)	113 (87 - 139)
Baseline Height Velocity (cm/year)		
Mean (SD)	3.9 (2.0)	3.9 (1.7)
Median (min - max)	3.8 (0.2 - 12.9)	4.3 (-0.9 - 6.2)
Missing	11	2
Baseline Height SDS		
Mean (SD)	-2.9 (0.8)	-3.0 (0.9)
Median (min - max)	-2.7 (-6.8 - -1.4)	-2.7 (-5.6 - -1.1)
Baseline BMI SDS		
Mean (SD)	-0.3 (0.9)	-0.1 (1.1)
Median (min - max)	-0.2 (-2.4 - 1.7)	-0.1 (-2.3 - 1.9)
Average Parental Height SDS		
Mean (SD)	-0.6 (0.8)	-0.4 (0.8)
Median (min - max)	-0.5 (-2.9 - 1.5)	-0.5 (-1.8 - 1.6)
Delta Mid-Parental Height SDS		
Mean (SD)	-2.3 (1.1)	-2.6 (1.3)
Median (min - max)	-2.2 (-7.2 - 0.1)	-2.1 (-6.5 - -0.8)
Peak GH Concentration (ng/ml)		
Mean (SD)	5.9 (2.8)	5.5 (3.0)
Median (min - max)	6.5 (0.2 - 10.0)	5.7 (0.2 - 10.0)
Log Peak GH Concentration		
Mean (SD)	1.6 (0.8)	1.4 (1.0)
Median (min - max)	1.9 (-1.4 - 2.3)	1.7 (-1.7 - 2.3)
Etiology Classification Code n (%)		

Treatment Group	Lonapeg.	Genotropin
II (%)	68 (65)	37 (66)
IO (%)	19 (18)	9 (16)
MPHD (%)	18 (17)	10 (18)

Abbreviations: Lonapeg-Lonapegsomatropin; GH-Growth Hormone; SDS-Standard Deviation Score; Delta Mid-Parental Height SDS – Father Height SDS Score – Mother Height SDS;II-Isolated Idiopathic; IO-Isolated Organic; MPHD-Multiple Pituitary Hormone Deficiencies*Age group was assessed at Visit 1.

3.2.4 Results and Conclusions

The primary endpoint of AHV demonstrated a statistically significant difference from Genotropin using the sponsor's pre-specified ANCOVA method (Table 4) and using my single imputation ANCOVA method (Table 5). Results from the single imputation ANCOVA assuming equal variances between arms were similar to results from single imputation ANCOVA assuming unequal variances. However, the treatment difference was less than one centimeter per year (cm/yr) for all analyses. This difference is more than 50% smaller in magnitude than the sponsor's pre-specified non-inferiority margin of 2 cm/yr.

The secondary endpoint of change in height SDS (Table 4) also demonstrated superiority to Genotropin using the sponsor's analysis method, though the magnitude of the treatment difference was also small.

Table 4: Primary and Secondary Endpoints – Sponsor's ANCOVA Results*

Endpoint*	Exp.	Ctrl.	Diff.	LCL	UCL	P-Val
Annualized Height Velocity (cm/yr.)	11.2	10.3	0.86	0.22	1.50	0.009
Change in Height SDS	1.10	0.96	0.14	0.03	0.26	0.015

*All endpoints are assessed at Week 52. No multiple testing procedure was used to control Type 1 error over primary and secondary endpoints; Abbreviations: cm/yr - centimeters per year; SDS-Standard Deviation Score; Exp.-Experimental Arm; Ctr.-Control Arm; Diff.-Treatment Difference;–LCL-Lower Confidence Limit; UCL Upper Confidence Limit ; P-Val-P-Value.

Table 5: Primary Endpoint – Reviewer's Results

Endpoint	Exp.	Ctrl.	Diff.	LCL	UCL	P-Val
Annualized Height Velocity (cm/yr.)*	11.1	10.3	0.80	0.13	1.47	0.02

Endpoint is assessed at Week 52. Abbreviations: cm/yr =centimeters per year; SDS-Standard Deviation Score; Exp.-Experimental Arm; Ctr.-Control Arm; Diff.-Treatment Difference; LCL-Lower Confidence Limit; UCL Upper Confidence Limit; P-Val-P-Value; Var.-Variance; *Reviewer's single imputation ANCOVA results- .

For the one-arm extension study CT-310EXT, the sponsor provided a bar graph (Figure 2) of change in height SDS with 13-week intervals up to Week 117, comparing patients who were

originally randomized to Lonapegsomatropin, to patients who were originally randomized to Genotropin and were then switched to Lonapegsomatropin for the extension study. The sponsor provided this as part of a response to a December 4, 2020 information request. This graph shows that both groups continued to show increases in change in height SDS over time during the extension study, and that the group originally randomized to Lonapegsomatropin continued to show a numerically higher change in height SDS at each time point compared to the group originally randomized to Genotropin, though the difference decreased over time. An ANCOVA with similar covariates and factors as used for the primary analysis (with the exception that the covariate of baseline height SDS –average SDS of parental height was not included) was used for this analysis. One limitation of the study is that only the non-missing values were included; the assessments of patients who had already discontinued treatment at each endpoint were not imputed for analysis. Also, datasets have not yet been submitted to verify these results. Of note, by week 104, only five of the 105 patients originally randomized to Lonapegsomatropin had discontinued treatment, and only three of the 56 patients originally randomized to Genotropin had discontinued treatment.

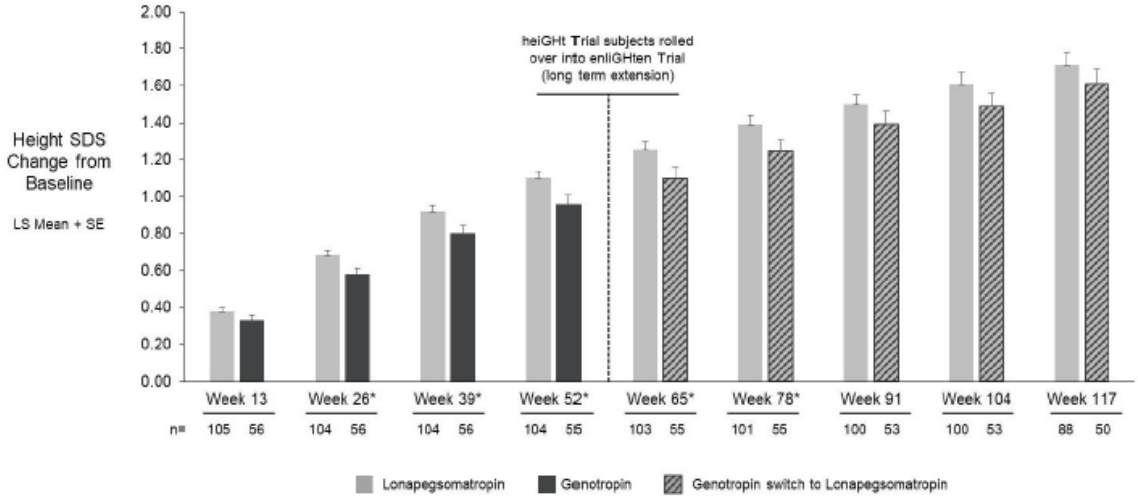


Figure 2: Change from Baseline in Height SDS by Visit
Reprinted from Figure 3 of sponsor's midcycle report

3.3 Evaluation of Safety

Only one patient, who was on the Lonapegsomatropin arm, discontinued treatment early before the Week 52 final assessment. This patient (Subject ID (b) (6)) had mild adverse events prior to treatment discontinuation, including respiratory tract disorders, bacterial infection, penile adhesion, and thyroid gland disorders. All these adverse events resolved except for the thyroid gland disorders. None of these adverse events were considered related to the study drug.

A total of 12 patients (11.4%) in the Lonapegsomatropin group and 10 patients (17.9%) on the Genotropin arm had adverse events that were considered related to treatment.
There was one SAE (serious adverse event) of appendicitis experienced by a patient (Subject ID (b) (6)) on the Genotropin arm. This SAE was considered not related to treatment. This patient was the patient who completed Genotropin treatment but had a missing final assessment. There was also a "mild" SAE of concussion experienced by a patient on the Genotropin arm. Both SAE's resolved and were considered not related to study drug. No TEAE led to discontinuation of study drug or death. There were also no reported deaths in this trial. Three TEAE's, one on the Genotropin arm and two on the Lonapegsomatropin arm, led to temporary dose reductions.

Safety results from the one-arm safety study CT-302 were similar in that there were few SAE's (only one patient had an SAE), only 4% of patients experienced a TEAE that was determined to be related to study drug, and no patient discontinued Lonapegsomatropin treatment due to an adverse event.

3.4 Benefit-Risk Assessment

The treatment difference of this drug compared to Genotropin was small – less than one cm/yr. However, superiority to another drug is not a requirement for approval. The AHV 11 cm/yr at 52 weeks is substantial compared to the baseline height velocity of 3.9 cm/yr. Only one patient on the Lonapegsomatropin discontinued treatment early, and the adverse event prior to discontinuation was mild and not considered related to the study drug. This drug also has the benefit of once weekly injections vs. daily injections. My conclusion is that the benefits outweigh the risks for the indicated pediatric population.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, Age, and Geographic Region

To assess the effect of Lonapegsomatropin compared to Genotropin within sex, age, region, and etiology, subgroup analyses were conducted for the primary endpoint using my preferred single imputation ANCOVA analysis defined in Section 3.2.2.2 These are shown in Table 6. Race subgroup analyses could not be conducted due to the sparsity of non-White race subgroups in the study. AHV was the outcome variable with baseline value, age, log peak concentration, and difference in parental height SDS as covariates, and treatment group and sex as factors.

Table 6: Treatment Difference in AHV, by Subgroup

Subgroup	Sample Size	Estimate (SE)	Lower 95%	Upper 95%
Overall	161	0.80 (0.34)	0.13	1.47
Female	29	0.59 (0.81)	-1.07	2.26

Male	132	0.74 (0.38)	-0.01	1.48
Age ≥ 6	122	0.43 (0.39)	-0.34	1.20
Age < 6	39	1.58 (0.65)	0.25	2.91
North Am.	27	0.23 (0.67)	-1.14	1.60
Europe	66	0.83 (0.46)	-0.09	1.75
ME North Afr.*	12	1.08 (1.15)	-1.51	3.67
Isolated Idiop.	105	0.58 (0.39)	-0.19	1.35
Isolated Organic**	28	1.09 (0.67)	-0.33	2.51
MPTH**	28	1.07 (1.21)	-1.43	3.56

Abbreviations: AHV- Annualized Height Velocity; SE – Standard Error;; ME-Middle East; North Afr.-North Africa; Idiop.-Idiopathic; MPTH – multiple pituitary hormone deficiencies* all the patients in this region are male; ** sex not included in model due to sparsity;

In the frequentist subgroup analysis shown in Table 6, there are random highs and random lows in sample estimates of subgroup treatment effects due to small sample size and large variability for some subgroups. Therefore, we also derived shrinkage estimates of subgroup treatment effects using a Bayesian hierarchical model based on summary sample estimates in Table 6. The total variability in the sample estimates is the sum of the within-subgroup variability of the sample estimator and the across-subgroup variability in the underlying/true parameter values. A shrinkage estimate of the subgroup treatment effect, which borrows information from the other subgroups while estimating the treatment effect for a specific subgroup, is a "weighted" average of the sample estimate and the overall estimate. The weights are based on the ratio of the between-subgroup variability to the within-subgroup variability. The greater the ratio, the smaller the weight on the overall estimate (the less the shrinkage). We used the same flat prior distribution to derive shrinkage estimates for all subgroups. The Bayesian hierarchical model assumptions are:

For $i = 1, 2\ldots$, Y_i represents the observed sample estimate of treatment effect in a subgroup level i, assume $Y_i \sim N(\mu_i, \sigma_i^2)$ where

- σ_i^2 are the observed variance for the subgroup sample estimates
- $\mu_i \sim N(\mu, \tau^2)$
- $\mu \sim N(0, 18^2)$, $1/\tau^2 \sim$ Gamma(0.001, 0.001)

The results of the sample and shrinkage estimates for treatment differences in subgroups are presented in Figure 3. The subgroup with the largest treatment difference, for both the frequentist and shrinkage analyses, was the Age<6 subgroup (1.58 cm/yr for the frequentist analysis, and

1.11 cm/yr for the shrinkage analysis). The subgroup with the smallest treatment difference for both the frequentist and the shrinkage analysis, was the North American subgroup (0.23 cm/yr for the frequentist analysis, and 0.59 cm/yr for the shrinkage analysis). However the shrinkage analysis estimate for the Age >=6 subgroup (0.60 cm/yr) was very close to that of the North American subgroup. Treatment effects were generally consistent across subgroups.

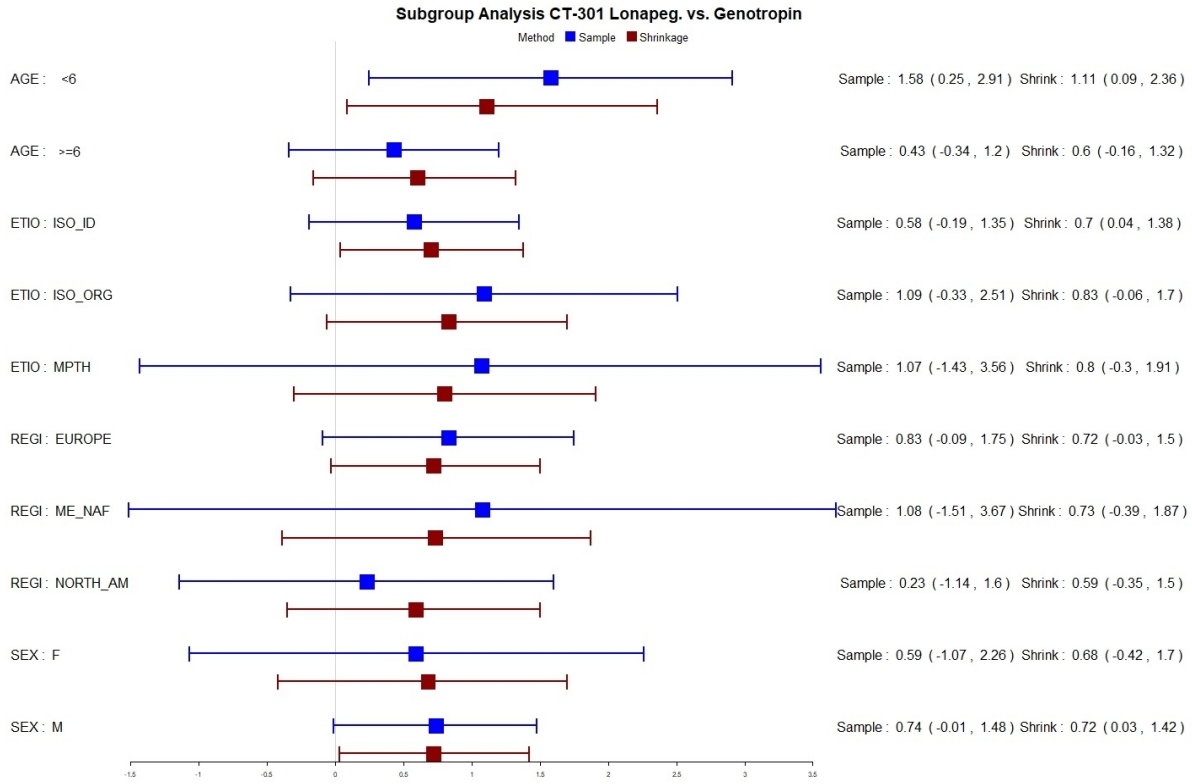


Figure 3:Forest Plot Comparing Frequentist Subgroup Analysis to Bayesian Shrinkage Analysis
Abbreviations: ME_NAF-Middle East/North Africa region; Etio-etiology; II-Isolated Idiopathic; IO-Isolated Organic; MPHD-Multiple Pituitary Hormone Deficiencies

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

The following are some potential statistical issues identified during the review.

- This primary endpoint is statistically significant in favor of Lonapegsomatropin in Study CT-301. However, the treatment difference is small, and a determination of clinical meaningfulness is needed.

(b) (4)

5.2 Collective Evidence

The primary analysis of the primary endpoint of AHV demonstrated statistical significance in favor of Lonapegsomatropin in comparison to Genotropin in Study CT-301. However, the treatment difference was more than 50% smaller in magnitude than the non-inferiority margin, (b) (4) The extension study for CT-301, a single arm study, provided some evidence that change in height SDS continues to improve well after the one-year efficacy endpoint. The missing data and discontinuation rates were very low, and no major safety issues have been identified. (b) (4) there is sufficiently strong evidence that Lonapegsomatropin is efficacious in the proposed indication and non-inferior to Genotropin.

5.3 Conclusions and Recommendations

I recommend that this drug be approved for the proposed indication. (b) (4)

5.4 Labeling Recommendations

The secondary endpoint of Change in Height SDS was not included in the hierarchical testing procedure. However, it correlates with the primary endpoint and may still provide useful information to patients and prescribers.

APPENDICES

Table 2. Analysis Window

VISIT	Week	Target Study Day	Study Day Window
Visit 1	Week 1	1	<=1
Visit 2	Week 13	90	2, 135
Visit 3	Week 26	181	136, 226
Visit 4	Week 39	272	227, 317
Visit 5	Week 52	363	318, 408
Visit 6	Week 65	454	409, 499
Visit 7	Week 78	545	500, 590
Visit 8	Week 91	636	591, 681
Visit 9	Week 104	727	682, 772
Visit x	Every 13 Weeks 13(x-1)	7*13*(x-1)-1 if x>1	Floor (7*13* (x-1.5)), Floor (7*13* (x-0.5))-1

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

ALEXANDER CAMBON
02/19/2021 01:46:41 PM

FENG LI
02/19/2021 01:54:45 PM

MARK D ROTHMANN
02/19/2021 02:07:26 PM
I concur